

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/402,093C
Source: 1FW/6
Date Processed by STIC: 7/8/05

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 07/08/2005

PATENT APPLICATION: US/09/402,093C

TIME: 11:47:16

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\07082005\I402093C.raw

4 <110> APPLICANT: Ohsuye, Kazuhiro
 5 Yabuta, Masayuki
 6 Suzuki, Yuji
 8 <120> TITLE OF INVENTION: Process for Producing Peptides Using a
 9 Helper Peptide
 11 <130> FILE REFERENCE: 001560-373
 13 <140> CURRENT APPLICATION NUMBER: US 09/402,093C
 14 <141> CURRENT FILING DATE: 1999-09-29
 16 <150> PRIOR APPLICATION NUMBER: PCT/JP99/00406
 17 <151> PRIOR FILING DATE: 1999-01-29
 19 <150> PRIOR APPLICATION NUMBER: JP 10-32272
 20 <151> PRIOR FILING DATE: 1998-01-30
 22 <160> NUMBER OF SEQ ID NOS: 70
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 4
 28 <212> TYPE: PRT
 29 <213> ORGANISM: Artificial Sequence
 31 <220> FEATURE:
 32 <223> OTHER INFORMATION: Amino acid sequence adjacent to a site cleaved by
 33 enterokinase
 35 <400> SEQUENCE: 1
 36 Asp Asp Asp Lys
 37 1
 40 <210> SEQ ID NO: 2
 41 <211> LENGTH: 4
 42 <212> TYPE: PRT
 43 <213> ORGANISM: Artificial Sequence
 45 <220> FEATURE:
 46 <223> OTHER INFORMATION: Amino acid sequence adjacent to a site cleaved by
 47 blood coagulation Factor Xa
 49 <400> SEQUENCE: 2
 50 Ile Glu Gly Arg
 51 1
 54 <210> SEQ ID NO: 3
 55 <211> LENGTH: 7
 56 <212> TYPE: PRT
 57 <213> ORGANISM: Artificial Sequence
 59 <220> FEATURE:
 60 <223> OTHER INFORMATION: Amino acid sequence containing a site cleaved by
 61 renin
 63 <400> SEQUENCE: 3
 64 Pro Phe His Leu Leu Val Tyr

P.6

RAW SEQUENCE LISTING

DATE: 07/08/2005

PATENT APPLICATION: US/09/402,093C

TIME: 11:47:16

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\07082005\I402093C.raw

```

65 1 5
68 <210> SEQ ID NO: 4
69 <211> LENGTH: 6
70 <212> TYPE: PRT
71 <213> ORGANISM: Artificial Sequence
73 <220> FEATURE:
74 <223> OTHER INFORMATION: Amino acid sequence of helper peptide
76 <400> SEQUENCE: 4
77 Val Asp Asp Asp Asp Lys
78 1 5
81 <210> SEQ ID NO: 5
82 <211> LENGTH: 6
83 <212> TYPE: PRT
84 <213> ORGANISM: Artificial Sequence
86 <220> FEATURE:
87 <223> OTHER INFORMATION: Amino acid sequence of helper peptide
89 <400> SEQUENCE: 5
90 Gly Cys His His His His
91 1 5
94 <210> SEQ ID NO: 6
95 <211> LENGTH: 10
96 <212> TYPE: PRT
97 <213> ORGANISM: Artificial Sequence
99 <220> FEATURE:
100 <223> OTHER INFORMATION: Amino acid sequence containing a chemically
101 cleaved site
103 <400> SEQUENCE: 6
104 Pro Gly Gly Arg Pro Ser Arg His Lys Arg
105 1 5 10
108 <210> SEQ ID NO: 7
109 <211> LENGTH: 10
110 <212> TYPE: PRT
111 <213> ORGANISM: Artificial Sequence
113 <220> FEATURE:
114 <223> OTHER INFORMATION: Amino acid sequence of helper peptide
116 <400> SEQUENCE: 7
117 His Arg His Lys Arg Ser His His His His
118 1 5 10
121 <210> SEQ ID NO: 8
122 <211> LENGTH: 5
123 <212> TYPE: PRT
124 <213> ORGANISM: Artificial Sequence
126 <220> FEATURE:
127 <223> OTHER INFORMATION: Amino acid sequence containing a site cleaved by
128 Kex2 Protease
130 <400> SEQUENCE: 8
131 Ser Asp His Lys Arg
132 1 5
135 <210> SEQ ID NO: 9

```

RAW SEQUENCE LISTING

DATE: 07/08/2005

PATENT APPLICATION: US/09/402,093C

TIME: 11:47:16

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\07082005\I402093C.raw

```

136 <211> LENGTH: 23
137 <212> TYPE: PRT
138 <213> ORGANISM: Artificial Sequence
140 <220> FEATURE:
141 <223> OTHER INFORMATION: Amino acid squence containing a position cleaved
142     by OmpT
144 <400> SEQUENCE: 9
145 Gln Met His Gly Tyr Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His
146 1           5           10           15
147 Arg Trp Gly Arg Ser Gly Ser
148     20
151 <210> SEQ ID NO: 10
152 <211> LENGTH: 20
153 <212> TYPE: PRT
154 <213> ORGANISM: Artificial Sequence
156 <220> FEATURE:
157 <223> OTHER INFORMATION: Amino acid sequence containing a position cleaved
158     by OmpT
160 <400> SEQUENCE: 10
161 Gln Met His Gly Tyr Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His
162 1           5           10           15
163 Gly Ser Gly Ser
164     20
167 <210> SEQ ID NO: 11
168 <211> LENGTH: 69
169 <212> TYPE: DNA
170 <213> ORGANISM: Artificial Sequence
172 <220> FEATURE:
173 <223> OTHER INFORMATION: Nucleotide sequence coding for an amino acid
174     sequence containing a site cleaved by OmpT
176 <400> SEQUENCE: 11
177 cagatgcatg gttatgacgc ggagctccgg ctgtatcgcc gtcatacccg gtgggggtcgt 60
178 tccggatcc                                     69
180 <210> SEQ ID NO: 12
181 <211> LENGTH: 23
182 <212> TYPE: PRT
183 <213> ORGANISM: Artificial Sequence
185 <220> FEATURE:
186 <223> OTHER INFORMATION: Amino acid sequence containing a site cleaved by
187     OmpT
189 <400> SEQUENCE: 12
190 Gln Met His Gly Tyr Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His
191 1           5           10           15
192 Arg Trp Gly Arg Ser Gly Ser
193     20
196 <210> SEQ ID NO: 13
197 <211> LENGTH: 47
198 <212> TYPE: DNA
199 <213> ORGANISM: Artificial Sequence

```

RAW SEQUENCE LISTING

DATE: 07/08/2005

PATENT APPLICATION: US/09/402,093C

TIME: 11:47:16

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\07082005\I402093C.raw

201 <220> FEATURE:
 202 <223> OTHER INFORMATION: Nucleotide sequence coding for an amino acid
 203 sequence containing a site cleaved by OmpT
 205 <400> SEQUENCE: 13
 206 tggttatgac gcgagactcc gcctgtatcg ccgtcatcac ggttcg 47
 208 <210> SEQ ID NO: 14
 209 <211> LENGTH: 55
 210 <212> TYPE: DNA
 211 <213> ORGANISM: Artificial Sequence
 213 <220> FEATURE:
 214 <223> OTHER INFORMATION: Nucleotide sequence coding for an amino acid
 215 sequence containing a site cleaved by OmpT
 217 <400> SEQUENCE: 14
 218 gatccggaac cgtgatgacg gcgatacagg cggagctccg cgtcataacc atgca 55
 220 <210> SEQ ID NO: 15
 221 <211> LENGTH: 24
 222 <212> TYPE: DNA
 223 <213> ORGANISM: Artificial Sequence
 225 <220> FEATURE:
 226 <223> OTHER INFORMATION: Primer
 228 <400> SEQUENCE: 15
 229 gactcagatc ttctgagggc cgat 24
 231 <210> SEQ ID NO: 16
 232 <211> LENGTH: 36
 233 <212> TYPE: DNA
 234 <213> ORGANISM: Artificial Sequence
 236 <220> FEATURE:
 237 <223> OTHER INFORMATION: Primer
 239 <400> SEQUENCE: 16
 240 aaaggtacct tccgcatgcc gcggatgtcg agaagg 36
 242 <210> SEQ ID NO: 17
 243 <211> LENGTH: 20
 244 <212> TYPE: DNA
 245 <213> ORGANISM: Artificial Sequence
 247 <220> FEATURE:
 248 <223> OTHER INFORMATION: Primer
 250 <400> SEQUENCE: 17
 251 aggccaggaa ccgtaaaaag 20
 253 <210> SEQ ID NO: 18
 254 <211> LENGTH: 29
 255 <212> TYPE: DNA
 256 <213> ORGANISM: Artificial Sequence
 258 <220> FEATURE:
 259 <223> OTHER INFORMATION: Primer
 261 <400> SEQUENCE: 18
 262 aaaatgcatc gcatcgtaac cgtgcatct 29
 264 <210> SEQ ID NO: 19
 265 <211> LENGTH: 627
 266 <212> TYPE: DNA

RAW SEQUENCE LISTING

DATE: 07/08/2005

PATENT APPLICATION: US/09/402,093C

TIME: 11:47:16

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\07082005\I402093C.raw

```

267 <213> ORGANISM: Artificial Sequence
269 <220> FEATURE:
270 <223> OTHER INFORMATION: Nucleotide sequence coding for a fusion protein
271     comprising GLP-1, helper peptide and
272     beta-galactosidase protective peptide
W--> 274 <221> NAME/KEY: CDS
275 <222> LOCATION: (82)...(543)
W--> 277 <400> 19
278 cccaggcttt acactttatg cttccggctc gtatgttggtg tggaattgtg agcggataac 60
279 aatttcacac aggaaacagc t atg acc atg att acg gat tca ctg gcc gtc 111
280                               Met Thr Met Ile Thr Asp Ser Leu Ala Val
281                               1           5           10
283 gtt tta caa cgt aaa gac tgg gat aac cct ggc gtt acc caa ctt aat 159
284 Val Leu Gln Arg Lys Asp Trp Asp Asn Pro Gly Val Thr Gln Leu Asn
285                               15           20           25
287 cgc ctt gca gca cat ccc cct ttc gcc agc tgg cgt aat agc gac gac 207
288 Arg Leu Ala Ala His Pro Pro Phe Ala Ser Trp Arg Asn Ser Asp Asp
289                               30           35           40
291 gcc cgc acc gat cgc cct tcc caa cag ttg cgc agc ctg aat ggc gaa 255
292 Ala Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu
293                               45           50           55
295 tgg cgc ttt gcc tgg ttt ccg gca cca gaa gcg gtg ccg gca agc ttg 303
296 Trp Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Ala Ser Leu
297                               60           65           70
299 ctg gag tca gat ctt cct gag gcc gat act gtc gtc gtc ccc tca aac 351
300 Leu Glu Ser Asp Leu Pro Glu Ala Asp Thr Val Val Val Pro Ser Asn
301 75           80           85           90
303 tgg cag atg cac ggt tac gat gcg atg cat ggt tat gac gcg gag ctg 399
304 Trp Gln Met His Gly Tyr Asp Ala Met His Gly Tyr Asp Ala Glu Leu
305                               95           100          105
307 cgc ctg tat cgc cgt cat cac ggt tcc gga tcc cct tct cga cat ccg 447
308 Arg Leu Tyr Arg Arg His His Gly Ser Gly Ser Pro Ser Arg His Pro
309                               110          115          120
311 cgg cat gcg gaa ggt acc ttt acc agc gat gtg agc tcg tat ctg gaa 495
312 Arg His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu
313                               125          130          135
315 ggt cag gcg gca aaa gaa ttc atc gcg tgg ctg gtg aaa ggc cgt ggt 543
316 Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
317                               140          145          150
319 taagtcgaca gcccgcttaa tgagcgggct tttttttctc ggaattaatt ctcatgtttg 603
320 acagcttatc atcgataagc tttta                                     627
322 <210> SEQ ID NO: 20
323 <211> LENGTH: 154
324 <212> TYPE: PRT
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: Amino acid sequence of a fusion protein comprising
329     GLP-1, helper peptide and beta-galactosidase
330     protective peptide

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/402,093C

DATE: 07/08/2005
TIME: 11:47:17

Input Set : A:\Sequence Listing.txt
Output Set: N:\CRF4\07082005\I402093C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:25; Xaa Pos. 6
Seq#:41; Xaa Pos. 2
Seq#:42; Xaa Pos. 2
Seq#:49; Xaa Pos. 2
Seq#:50; Xaa Pos. 2
Seq#:51; Xaa Pos. 2
Seq#:52; Xaa Pos. 2
Seq#:53; Xaa Pos. 2
Seq#:54; Xaa Pos. 2
Seq#:61; Xaa Pos. 2
Seq#:62; Xaa Pos. 2
Seq#:63; Xaa Pos. 2
Seq#:64; Xaa Pos. 2
Seq#:65; Xaa Pos. 2
Seq#:66; Xaa Pos. 2
Seq#:69; Xaa Pos. 2
Seq#:70; Xaa Pos. 2

VERIFICATION SUMMARY

DATE: 07/08/2005

PATENT APPLICATION: US/09/402,093C

TIME: 11:47:17

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\07082005\I402093C.raw

L:274 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:277 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:19
L:489 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:493 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:25
L:494 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0
L:523 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:527 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:27
L:602 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:606 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:32
L:621 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:625 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:33
L:640 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:644 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:34
L:757 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:761 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:41
L:762 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:0
L:776 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:780 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:42
L:784 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:42
L:785 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42 after pos.:0
L:814 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:818 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:44
L:848 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:852 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:46
L:882 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:886 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:48
L:901 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:905 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:49
L:906 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:0
L:920 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:924 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:50
L:928 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:50
L:929 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0
L:943 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:947 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:51
L:948 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:0
L:962 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:966 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:52
L:970 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:52
L:971 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:0
L:985 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:989 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:53
L:990 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53 after pos.:0
L:1004 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1008 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:54
L:1012 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:54
L:1013 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:0
L:1042 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!

VERIFICATION SUMMARY

DATE: 07/08/2005

PATENT APPLICATION: US/09/402,093C

TIME: 11:47:17

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\07082005\I402093C.raw

L:1046 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:56
L:1076 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1080 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:58
L:1110 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1114 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:60
L:1129 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1133 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:61
L:1134 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61 after pos.:0
L:1148 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1152 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:62
L:1156 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:62
L:1157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:62 after pos.:0
L:1171 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1175 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:63
L:1176 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63 after pos.:0
L:1190 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1194 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:64
L:1198 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:64
L:1199 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:64 after pos.:0
L:1213 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1217 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:65
L:1218 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65 after pos.:0
L:1232 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1236 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:66
L:1240 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:66
L:1241 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66 after pos.:0
L:1270 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1274 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:68
L:1289 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1293 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:69
L:1294 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:69 after pos.:0
L:1308 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:1312 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:70
L:1316 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:70
L:1317 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70 after pos.:0